# Package ‘DAMOCLES’

March 5, 2015

<table>
<thead>
<tr>
<th>Type</th>
<th>Package</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Dynamic Assembly Model of Colonization, Local Extinction and Speciation</td>
</tr>
<tr>
<td>Version</td>
<td>1.1</td>
</tr>
<tr>
<td>Date</td>
<td>2015-03-05</td>
</tr>
<tr>
<td>Depends</td>
<td>R (&gt;= 3.0), geiger, caper, ape, deSolve, matrixStats</td>
</tr>
<tr>
<td>Imports</td>
<td>expm, picante</td>
</tr>
<tr>
<td>Author</td>
<td>Rampal S. Etienne &amp; Alex L. Pigot</td>
</tr>
<tr>
<td>Maintainer</td>
<td>Rampal S. Etienne <a href="mailto:r.s.etienne@rug.nl">r.s.etienne@rug.nl</a></td>
</tr>
<tr>
<td>License</td>
<td>GPL-2</td>
</tr>
<tr>
<td>Description</td>
<td>Simulates and computes (maximum) likelihood of a dynamical model of community assembly that takes into account phylogenetic history.</td>
</tr>
<tr>
<td>NeedsCompilation</td>
<td>no</td>
</tr>
<tr>
<td>Repository</td>
<td>CRAN</td>
</tr>
<tr>
<td>Date/Publication</td>
<td>2015-03-05 17:36:50</td>
</tr>
</tbody>
</table>

## R topics documented:

- DAMOCLES-package .................................................. 2
- DAMOCLES_bootstrap .............................................. 3
- DAMOCLES_loglik .................................................... 6
- DAMOCLES_ML .......................................................... 8
- DAMOCLES_sim ....................................................... 10
- NWPrimates_data ................................................... 12

Index 13
Description

Simulation and likelihood methods for a dynamical community assembly that accounts for phylogenetic history

New in version 1.1:
- Function added to do bootstrap

Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>DAMOCLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>1.1</td>
</tr>
<tr>
<td>Date:</td>
<td>2015-03-05</td>
</tr>
<tr>
<td>License:</td>
<td>GPL 2.0</td>
</tr>
</tbody>
</table>

DAMOCLES_loglik computes the likelihood of presence-absence data in a local community given a set of parameters and a phylogeny under an immigration-extinction model

DAMOCLES_ML finds the parameters that maximizes the likelihood computed by DAMOCLES_loglik.

DAMOCLES_sim simulates presence-absence data for a given phylogeny

DAMOCLES_bootstrap computes the maximum likelihood estimates of colonisation and local extinction rate for a given phylogeny and presence-absence data under the DAMOCLES model. These rate estimates are used to simulate null communities under the DAMOCLES model. Standardized effect size of mean nearest taxon distance (mntd), mean phylogentic distance (mpd) and loglikelihood are calculated. For comparison, standardised effect sizes are also calculated relative to a "Random-Draw" null model i.e. presence absence randomised across tips.

Author(s)

Rampal S. Etienne & Alex L. Pigot
Maintainer: Rampal S. Etienne (r.s.etienne@rug.nl)

References

**DAMOCLES_bootstrap**

**Examples**

DAMOCLES_ML()

---

**Phylogenetic community structure hypothesis test**

**Description**

This function computes the maximum likelihood estimates of colonisation and local extinction rate for a given phylogeny and presence-absence data under the DAMOCLES model. These rate estimates are used to simulate null communities under the DAMOCLES model. Standardized effect size of mean nearest taxon distance (mntd), mean phylogenetic distance (mpd) and loglikelihood are calculated. For comparison, standardised effect sizes are also calculated relative to a "Random-Draw" null model i.e. presence absence randomised across tips.

**Usage**

DAMOCLES_bootstrap(
  phy = rcoal(10),
  pa = matrix(c(phy$tip.label, sample(c(0,1),Ntip(phy),replace = T)),
    nrow = Ntip(phy),ncol = 2),
  initparsopt = c(0.1,0.1),
  idparsopt = 1:length(initparsopt),
  parsfix = 0,
  idparsfix = (1:3)[-idparsopt],
  pars2 = c(1E-3,1E-4,1E-5,1000),
  pchoice = 0,
  runs = 999,
  estimate_pars = FALSE,
  conf.int = 0.95
)

**Arguments**

- `phy` phylogeny in phylo format
- `pa` presence-absence table.
  The first column contains the labels of the species (corresponding to the tip labels in the phylogeny).
  The second column contains the presence (1) or absence (0) of species in the local community.
- `initparsopt` The initial values of the parameters that must be optimized.
- `idparsopt` The ids of the parameters that must be optimized, e.g. 1:2 for extinction rate, and offset of immigration rate. The ids are defined as follows:
  - id == 1 corresponds to mu (extinction rate)
  - id == 2 corresponds to gamma_0 (offset of immigration rate)
parsfix: The values of the parameters that should not be optimized. See idparsfix.

idparsfix: The ids of the parameters that should not be optimized, e.g. c(1) if mu should not be optimized, but only gamma_0. In that case idparsopt must be c(2). The default is to fix the parameters not specified in idparsopt.

pars2: Vector of settings:
- pars2[1] sets the relative tolerance in the parameters
- pars2[2] sets the relative tolerance in the function
- pars2[3] sets the absolute tolerance in the parameters
- pars2[4] sets the maximum number of iterations

pchoice: sets which p-value to optimize and with which root state to simulate (default pchoice = 0)
- pchoice == 0 correspond to optimizing sum of p_0f + p_1f, and simulating with an equal number of root states being 0 or 1
- pchoice == 1 correspond to optimizing p_0f, and simulating with root state being 0
- pchoice == 2 correspond to optimizing p_1f, and simulating with root state being 1

runs: the number null communities to generate.

estimate_pars: Whether to estimate parameters on the simulated datasets (default = FALSE).

conf.int: The width of the confidence intervals calculated on bootstrapped parameter estimates

**Details**

The output is a list of two dataframes. The first dataframe, summary_table, contains the summary results. The second dataframe, null_community_data, contains descriptive statistics for each null community.

**Value**

**summary_table**
- mu gives the maximum likelihood estimate of mu and confidence intervals in brackets if estimate_pars = TRUE
- gamma_0 gives the maximum likelihood estimate of gamma_0 and confidence intervals in brackets if bootstrap=TRUE
- loglik gives the maximum loglikelihood
- df gives the number of estimated parameters, i.e. degrees of freedom
- conv gives a message on convergence of optimization; conv = 0 means convergence
- n.obs gives the number of species locally present in the observed community
- mntd.obs gives the MNTD of the observed community
- mpd.obs gives the MPD of the observed community
- runs gives the number of null communities simulated
- mntd.mean: mean of MNTD from null communities generated by a "Random Draw" model
- mntd.sd: standard deviation of MNTD from null communities generated by a "Random Draw" model
- mntd.obs.z.RD: standardized effect size of MNTD compared to null communities generated by a "Random Draw" model
null_community_data

run gives the simulation run root.state.print gives the state of the ancestral species in the local community assumed in the simulation, i.e. present (1) or absent (0) n gives the number of species locally present in the observed community n.RD gives the number of species locally present in the null community generated by a "Random Draw" model mntd.RD gives the MNTD of the null community generated by a "Random Draw" model mntd.mean.RD gives the mean of MNTD from null communities generated by a "Random Draw" model mntd.sd.RD gives the standard deviation of MNTD from null communities generated by a "Random Draw" model mntd.obs.rank.RD gives the rank of observed MNTD compared to null communities generated by a "Random Draw" model mntd.obs.q.RD quantile of observed MNTD vs. null communities (= mntd.obs.rank.RD /runs + 1) mntd.NobsNrankNrd rank of observed MNTD compared to null communities generated by a "Random Draw" model mntd.NobsNqNrd quantile of observed MNTD vs. null communities (= mntd.obs.rank.RD /runs + 1) mntd.NmeanNrd mean of MNTD from null communities generated by a "Random Draw" model mntd.NsdNrd standard deviation of MNTD from null communities generated by a "Random Draw" model mntd.NobsNzNrd standardized effect size of MNTD compared to null communities generated by a "Random Draw" model (= -1*(mntd.obs - mntd.mean.RD)/ mntd.sd.RD) mntd.NobsNrankNrd rank of observed MNTD compared to null communities generated by a "Random Draw" model mntd.NobsNqNrd quantile of observed MNTD vs. null communities (= mntd.obs.rank.RD /runs + 1) mpd.NmeanNrd mean of MPD from null communities generated by a "Random Draw" model mpd.NsdNrd standard deviation of MPD from null communities generated by a "Random Draw" model mpd.NobsNrankNrd rank of observed MPD compared to null communities generated by a "Random Draw" model mpd.NobsNqNrd quantile of observed MPD vs. null communities (= mpd.obs.rank.RD /runs + 1) n.mean.DAMOCLES mean number of species locally present in the null communities generated by DAMOCLES mntd.Nmean.DAMOCLES mean of MNTD from null communities generated by DAMOCLES mntd.Nsd.DAMOCLES standard deviation of MNTD from null communities generated by DAMOCLES mntd.Nobs.z.DAMOCLES standardized effect size of MNTD compared to null communities generated by DAMOCLES (= -1*(mntd.obs - mntd.mean.DAMOCLES)/ mntd.sd.DAMOCLES) mntd.Nobs.rank.DAMOCLES rank of observed MNTD compared to null communities generated by DAMOCLES mntd.Nobs.q.DAMOCLES quantile of observed MNTD vs. null communities (= mntd.obs.rank.DAMOCLES /runs + 1) mpd.Nmean.DAMOCLES mean of MPD from null communities generated by DAMOCLES mpd.Nsd.DAMOCLES standard deviation of MPD from null communities generated by DAMOCLES mpd.Nobs.z.DAMOCLES standardized effect size of MPD compared to null communities generated by DAMOCLES (= -1*(mpd.obs - mpd.mean.DAMOCLES)/ mpd.sd.DAMOCLES) mpd.Nobs.rank.DAMOCLES rank of observed MPD compared to null communities generated by DAMOCLES mpd.Nobs.q.DAMOCLES quantile of observed MPD vs. null communities (= mpd.obs.rank.DAMOCLES /runs + 1) loglik.Nmean.DAMOCLES mean of loglikelihoods from null communities generated by DAMOCLES loglik.Nsd.DAMOCLES standard deviation of loglikelihoods from null communities generated by DAMOCLES loglik.Nobs.z.DAMOCLES standardized effect size of loglikelihoods compared to null communities generated by DAMOCLES (= -1*(loglik.obs - loglik.mean.DAMOCLES)/ loglik.sd.DAMOCLES) loglik.Nobs.rank.DAMOCLES rank of observed loglikelihood compared to null communities generated by DAMOCLES loglik.Nobs.q.DAMOCLES quantile of observed loglikelihoods vs. null communities (= loglik.obs.rank.DAMOCLES /runs + 1) loglik.Nmean.DAMOCLES mean of loglikelihoods from null communities generated by DAMOCLES loglik.Nsd.DAMOCLES standard deviation of loglikelihoods from null communities generated by DAMOCLES loglik.Nobs.z.DAMOCLES standardized effect size of loglikelihoods compared to null communities generated by DAMOCLES (= -1*(loglik.obs - loglik.mean.DAMOCLES)/ loglik.sd.DAMOCLES) loglik.Nobs.rank.DAMOCLES rank of observed loglikelihood compared to null communities generated by DAMOCLES loglik.Nobs.q.DAMOCLES quantile of observed loglikelihoods vs. null communities (= loglik.obs.rank.DAMOCLES /runs + 1)
likelihood estimate of $\mu$ for the null community generated by DAMOCLES

$\gamma_0$. DAMOCLES gives the maximum likelihood estimate of $\gamma_0$ for
the null community generated by DAMOCLES

Author(s)
Rampal S. Etienne

References

See Also
DAMOCLES_ml, DAMOCLES_sim

Examples
```r
data(nwprimates_data)
out = DAMOCLES_bootstrap(
  phy = nwprimates_data[[1]],
  pa = nwprimates_data[[2]],
  initparsopt = c(0.01,1.8),
  idparsopt = c(1,2),
  pars2 = c(1E-3,1E-4,1E-5,1000),
  pchoice = 1,
  runs = 2,
  estimate_pars = TRUE,
  conf.int = 0.95)
```

DAMOCLES_loglik  Likelihood for DAMOCLES model

Description
Computes likelihood for the presence-absence data of species in a local community for a given phylogeny of species in the region.

Usage
```r
DAMOCLES_loglik(
  phy,
  pa,
  pars,
  pchoice = 0,
  edgeTList = compute_edgeTList(phy)
)```
**Arguments**

- **phy**  
  phylogeny in phylo format

- **pa**  
  presence-absence table with the first column the species labels and the second column the presence (1) or absence (0) of the species

- **pars**  
  Vector of model parameters:  
  - `pars[1]` corresponds to \( \mu \) (extinction rate in local community)  
  - `pars[2]` corresponds to \( \gamma_0 \) in formula \( \gamma(t) = \gamma_0/(1 + \gamma_1 t) \) where \( \gamma(t) \) is immigration rate into local community  
  - `pars[3]` corresponds to \( \gamma_1 \) in formula \( \gamma(t) = \gamma_0/(1 + \gamma_1 t) \) where \( \gamma(t) \) is immigration rate into local community

- **pchoice**  
  sets the p-value to optimize:  
  - `pchoice == 0` corresponds to the sum of \( p_{0f} + p_{1f} \)  
  - `pchoice == 1` corresponds to \( p_{0f} \)  
  - `pchoice == 2` corresponds to \( p_{1f} \)

- **edgeTList**  
  list of edge lengths that need to be succesively pruned; if not specified, it will computed using `compute_edgeTList`

**Value**

The loglikelihood

**Author(s)**

Rampal S. Etienne

**References**


**See Also**

`DAMOCLES_ML` `DAMOCLES_sim`

**Examples**

```r
#TEST IT WORKS  
phy = rcoal(100)  
pars = c(0.5,0.1,0.1)  
pa = rbinom(100,c(0,1),0.5)  
pa = matrix(c(phy$tip.label,pa),nrow = length(phy$tip.label),ncol = 2)

# - without a root edge  
loglik = DAMOCLES_loglik(phy,pa,pars)  
loglik

# - with a root edge  
phy$root.edge = 2
```
loglik = DAMOCLES_loglik(phy, pa, pars)
loglik

---

### DAMOCLES_ML

**Maximization of the loglikelihood under the DAMOCLES model**

**Description**

This function computes the maximum likelihood estimates of the parameters of the DAMOCLES model for a given phylogeny and presence-absence data. It also outputs the corresponding loglikelihood that can be used in model comparisons.

**Usage**

```
DAMOCLES_ML(
  phy = rcoal(10),
  pa = matrix(c(phy$tip.label, sample(c(0,1), Ntip(phy), replace = T)),
              nrow = Ntip(phy), ncol = 2),
  initparsopt = c(0.1, 0.1),
  idparsopt = 1:length(initparsopt),
  parsfix = 0,
  idparsfix = (1:3)[-idparsopt],
  pars2 = c(1E-3, 1E-4, 1E-5, 1000),
  pchoice = 0
)
```

**Arguments**

- **phy**: phylogeny in phylo format
- **pa**: presence-absence table.
  The first column contains the labels of the species (corresponding to the tip labels in the phylogeny).
  The second column contains the presence (1) or absence (0) of species in the local community.
- **initparsopt**: The initial values of the parameters that must be optimized
- **idparsopt**: The ids of the parameters that must be optimized, e.g. 1:2 for extinction rate, and offset of immigration rate. The ids are defined as follows:
  - id == 1 corresponds to mu (extinction rate)
  - id == 2 corresponds to gamma_0 (offset of immigration rate)
  - id == 3 corresponds to gamma_1 (parameter controlling decline in immigration rate with time)
- **parsfix**: The values of the parameters that should not be optimized. See idparsfix.
- **idparsfix**: The ids of the parameters that should not be optimized, e.g. c(1,3) if mu and gamma_1 should not be optimized, but only gamma_0. In that case idparsopt must be c(2). The default is to fix all parameters not specified in idparsopt.
pars2 Vector of settings:
pars2[1] sets the relative tolerance in the parameters
pars2[2] sets the relative tolerance in the function
pars2[3] sets the absolute tolerance in the parameters
pars2[4] sets the maximum number of iterations

pchoice sets which p-value to optimize (default 0)
pchoice == 0 correspond to the sum of p_0f + p_1f
pchoice == 1 correspond to p_0f
pchoice == 2 correspond to p_1f

Details
The output is a dataframe containing estimated parameters and maximum loglikelihood.

Value
mu gives the maximum likelihood estimate of mu
gamma_0 gives the maximum likelihood estimate of gamma_0
gamma_1 gives the maximum likelihood estimate of gamma_1
loglik gives the maximum loglikelihood
df gives the number of estimated parameters, i.e. degrees of freedom
conv gives a message on convergence of optimization; conv = 0 means convergence

Author(s)
Rampal S. Etienne

References

See Also
DAMOCLES_loglik DAMOCLES_sim

Examples
```r
data(NWPrimates_data)
out = DAMOCLES_ML(
  phy = NWPrimates_data[[1]],
  pa = NWPrimates_data[[2]],
  initparsopt = c(0, 0.1, 1.8),
  idparsopt = c(1, 2),
  parsfix = c(0),
  idparsfix = c(3),
)```
pars2 = c(1E-3,1E-4,1E-5,1000),
pchoice = 0)

DAMOCLES_sim
Simulating DAMOCLES

Description
Simulates DAMOCLES

Usage
DAMOCLES_sim(
  phy,
  gamma_0,
  gamma_td,
  mu,
  sigma,
  psiBranch,
  psiTrait,
  z,
  phi,
  traitOpt,
  br0,
  br_td,
  nTdim,
  root.state,
  root.trait.state,
  plotit = FALSE,
  keepExtinct = FALSE
)

Arguments
phy        phylogeny in phylo format
gamma_0    initial per lineage rate of immigration (gamma)
gamma_td   time dependency in gamma
mu          per lineage rate of local extinction
sigma      probability of local (i.e. in-situ) speciation
psiBranch  phylogenetic distance at which gamma is half gamma_0
psiTrait   trait distance at which gamma is half gamma_0
z          shape of increase in gamma with increasing trait or phylogenetic distance
phi        rate of decline in gamma with distance from trait optima
traitOpt   trait value at which gamma = gamma_0
**DAMOCLES_sim**

- **br0**  Brownian rate parameter
- **br_td**  rate of temporal decline in Brownian rate parameter
- **nTdim**  number of independent trait dimensions
- **root.state**  geographic state of ancestor i.e. present (1) or absent (0)
- **root.trait.state**  trait value of ancestor
- **plotit**  whether to plot the phylogeny and timing of immigration/local extinction events
- **keepExtinct**  whether to retain data for extinct lineages

**Value**

A list of two tables. The first table contains the following columns: The first column contains the vector of tip labels in the phylogeny. The last column contains the presence (1) or absence (0) of the species. The second table has dimensions d x N where d is the number of trait dimensions and N is the number of species. It contains the trait values.

**Author(s)**

Alex L. Pigot

**References**


**See Also**

[DAMOCLES_ML](#) [DAMOCLES_loglik](#)

**Examples**

```r
# create random phylogeny
phy <- rcoal(10)

# run DAMOCLES
out <- DAMOCLES_sim(
  phy,
  gamma_0 = 1.5,
  gamma_td = 0,
  mu = 0,
  sigma = 0,
  psiBranch = 0,
  psiTrait = 0,
  z = 10,
  phi = 0,
  traitOpt = 1,
  br0 = 0.1,
  br_td = -0.1,
  nTdim = 2,
)```
root.state = 1, 
root.trait.state = 0, 
plotit = FALSE, 
keepExtinct = FALSE 
)

#the output consists of a list 
patable = out[[1]] # the first element is the presence absence table 
traits = out[[2]] # this is a matrix of traits values 

#show presence/absence on the tree 
patable$col = rep("black",dim(patable)[1]) 
patable$col[which(patable$state == 1)] = "red" 
plot(phy,tip.col = patable$col)

NWPrimates_data

Dated phylogenetic tree of the New World Primates in nexus format and presence-absence matrix for species in Manu

Description

A list with two elements. 
. phy is a dated molecular phylogeny for 94 species of New World Primates extracted from the maximum likelihood tree (AUTOsoft dated) of Springer et al. (2012). 1 time unit = 100 million years. 
. pa is the presence-absence matrix of NW Primates in Manu from Solari et al. (2006). The first column indicate the species tip labels and the second column indicates presence (1) and absence (0).

Usage

data(NWPrimates_data)

Format

A list with two elements. The first element (phy) is the primate phylogeny in nexus format. The second element (pa) is the presence-absence matrix with 94 rows and 2 columns.

Source


See Also

DAMOCLES_sim, DAMOCLES_ML, DAMOCLES_loglik
Index

*Topic datasets
  NWPrimates_data, 12

*Topic models
  DAMOCLES-package, 2
  DAMOCLES_bootstrap, 3
  DAMOCLES_loglik, 6
  DAMOCLES_ML, 8
  DAMOCLES_sim, 10

DAMOCLES (DAMOCLES-package), 2
DAMOCLES-package, 2
DAMOCLES_bootstrap, 3
DAMOCLES_loglik, 6, 9, 11, 12
DAMOCLES_ML, 6, 7, 8, 11, 12
DAMOCLES_sim, 6, 7, 9, 10, 12

NWPrimates_data, 12